OPTIMIZED DEEP LEARNING TECHNIQUES FOR DETECTION AND CLASSIFICATION OF BRAIN TUMOR IN MRI IMAGES

A SYNOPSIS

Submitted by

S. POORNAM

(Reg. No. 9200104004)

in partial fulfillment for the award of the degree

of

DOCTOR OF PHILOSOPHY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SCHOOL OF COMPUTING

KALASALINGAM ACADEMY OF RESEARCH AND EDUCATION

(DEEMED TO BE UNIVERSITY)

ANAND NAGAR, KRISHNANKOIL – 626126

TAMILNADU, INDIA

NOVEMBER 2024

1. INTRODUCTION

The brain is the most important unit in the body of a person since it controls all other organ processes and facilitates making choices. It is regarded as the body's most important and crucial organ. An unusual development or clump of brain cells is called a "brain tumor." Brain tumors come in many different forms and can be classified as malignant (cancerous) or benign (noncancerous). Brain tumors originate in the brain after spreading from a different part of the body, primary brain tumors originate in the brain. The brain works under billions of cells and the highest mortality rate of tumors is reported for adults and children. The brain regulates diverse biological tasks such as decision-making, integration, organization, giving orders, and analyzing. In recent years, BT cases have been detected according to diverse scanning schemes like Computer Tomography (CT) and Magnetic Resonance Imaging (MRI) etc. The CT pictures are responsible for the detection of soft tissues and MRI assists in identifying the hard tissues in brain area. Generally, BT cases are categorized into meningioma, glioblastoma, and glioma the categorization varies based on the surrounding region, the form, and the dimensions of the tumors. In the past decades, machine learning (ML) models have acted as a means to increase efficacy of different challenges like diagnosis, identification, tissue segmentation, and classification. However, these ML methods are affected because of challenges of computational complexity, lack of data interpretability, data availability, lack of transparency, and overfitting.

This research work aims to explore and compare the effectiveness of machine learning and deep learning methods in classifying brain tumors based on medical imaging data. Specifically, the focus will be on analyzing magnetic resonance imaging (MRI) scans, which are commonly used for diagnosing brain tumors due to their superior soft tissue contrast and spatial resolution. By leveraging Deep Learning techniques, this research seeks to develop robust classification models capable of accurately distinguishing between different tumor types, including gliomas, meningiomas, and pitutary.

2. LITERATURE SURVEY

The brain is the most important unit in the body of a person since it controls all other organ processes and facilitates making choices. It is regarded as the body's most important and crucial organ. An unusual development or clump of brain cells is called a

"brain tumor." Brain tumors come in many different forms and can be classified as malignant (cancerous) or benign (noncancerous). Whereas secondary, or metastatic, brain tumors originate in the brain after spreading from a different part of the body, primary brain tumors originate in the brain.

Ahamed et al. [1] using automated techniques, division of brain tumors has emerged as a common task in medical imaging to identify the location, size, and shape of the tumor. Many researchers have employed diverse machine learning and deep learning techniques to ascertain the most effective resolution through the application of convolutional technology.

Atmakuri et al. [2] Suggested to handling complex patterns and supervised categorization, deep learning has proven to be the most effective approach. In the simulation environment, a completely automated deep learning method can be created using the current toolboxes. The goal of the project is to classify photos of brains with meningioma, glioma, and pituitary tumors using this machine learning technique. Standard metrics are employed in the analysis, which is done in a Matlab environment during the simulation. The accuracy range displayed in the confusion matrix is 89% to 95%. In order to attain optimal outcomes, the Convolutional Neural Network (CNN) and Artificial Neural Network (ANN) models are contrasted with the obtained data.

Turk et al. [3] implemented a proposed system in three stages. In the first stage, it was determined whether there was a tumor in the MR images (Binary Approach). In the second stage, different tumor types (Normal, Glioma Tumor, Meningioma Tumor, Pituitary Tumor) were detected from MR images (Multi-class Approach). In the last stage, CAMs of each tumor group were created as an alternative tool to facilitate the work of specialists in tumor detection.

Talukder et al. [4] suggested to use four transfer learning architectures—InceptionV3, VGG19, DenseNet121, and Mobile Net—are used in this research. To validate the models, we employed a dataset comprising information from three benchmark databases: figshare, SARTAJ, and Br35H. There are four classes in these databases: glioma, meningioma, pituitary, and no tumor. The purpose of image augmentation is to balance the classes. The accuracy of 99.60% in the experimental findings shows that Mobile Net works better than competing techniques.

Patil et al. [5] Xception, ResNet50V2, InceptionResNetV2, and DenseNet201 accuracy scores were 99.40%, 99.68%, 99.36%, and 98.72%, respectively, based on our studies using the 3,064-image Figshare MRI brain tumor

dataset. Our results show that ResNet50V2 outperforms other models, achieving the greatest accuracy rate of 99.68% on the Figshare MRI brain tumor dataset. Thus, the capacity of our suggested model to quickly and reliably classify brain cancers can help neurologists and physicians make timely and accurate diagnoses for patients with brain tumors.

Tahosin et al. [6] suggested a method involved designing a shallow convolutional neural network (SCNN) and a VGG16 network using an MRI image of the T1C modality first, and then analyzing accuracy and loss. The collected features from both deep learning models were merged to increase the classification accuracy of three different types of tumors, hence improving the model's performance in terms of accuracy and loss information. The results obtained from the ensemble deep convolutional neural network model (EDCNN) demonstrated that the combination of deep learning models attempts to address the issue of overfitting of the model for unbalanced datasets while also improving the accuracy of the multiclass classification task. The suggested model aims to provide a 97.77% classification accuracy. Additionally, the suggested framework outperforms other state-of-the-art frameworks in terms of performance.

Turk et al. [7] implemented the three phases to the proposed system's. The first step was determining if the MR images showed evidence of a malignancy (Binary Approach). Using MR images and a multi-class approach, several tumor types (normal, glioma, meningioma, and pituitary) were identified in the second stage. To help professionals in tumor detection in their work, CAMs of each tumor category were developed in the final stage. Based on the ResNet50, InceptionV3, and Mobile Net architectures, the overall accuracy of the binary technique was calculated to be 100%, while on the VGG19 architecture, it was calculated to be 99.71%. Furthermore, the multi-class technique yielded accuracy values of 96.45% with ResNet50, 93.40% with VGG19, 85.03% with InceptionV3, and 89.34% with Mobile Net architectures Architecture VGG19.

Puustinen et al. [8] suggested to use publicly accessible dataset of 20 patients with newly diagnosed glioblastoma has been used to assess the method's performance and compare it with numerous cutting-edge methods. The approach achieves the greatest values of 0.96, 0.94, 0.96, 0.94, and 0.90, respectively, for accuracy, precision, specificity, F1-score, sensitivity, and MCC, according to the experimental data. Additionally, the method's stability and resilience under various noise levels and image

resolutions have been demonstrated. The suggested method has the potential to be used for additional medical image processing jobs and provides a quick, reliable, and accurate identification of brain tumors.

Raghuram and Bhukya Hanumanthu [9] proposed efficient brain image recognition in the internet of medical things to identify and categorize harmful cells in brain image recognition. The input brain image is beheaded in order to extract the brain area effectively. Then, using the extracted features as a guide, the suggested SDNN recognition classifies the brain images as normal or abnormal, extracting the dominant features from the skull-stripped images in the process.

Chatterjee et al. [10] included Data augmentation, learning transfer, and image acquisition sequences—T1-weighted images (T1WI), T1-weighted post-gadolinium (T1-Gd), and Fluid-Attenuated Inversion Recovery (FLAIR) in the development of brain tumor detection algorithm. According to the findings, accuracy was raised by up to 6% by employing learning transfer and data augmentation, with a p-value below the significance level of 0.05. The FLAIR sequence proved to be the most effective in terms of detection as well. On the other hand, our suggested model worked best in terms of training time, requiring around half of the second quickest network's time.

Anantharajan et al. [11] suggested a novel MRI brain tumour detection method based on DL and ML. Initially the MRI images are collected and preprocessed using Adaptive Contrast Enhancement Algorithm (ACEA) and median filter. Ranjbarzadeh et al. [12] suggested an automatic and robust brain tumor segmentation framework using four MRI sequence images, an optimized Convolutional Neural Network (CNN) and weight and bias values of the CNN model are adjusted using an Improved Chimp Optimization Algorithm (IChOA).

Lamba et al. [13] deployed a visual geometry group with 16 layers following transfer learning to help minimize the medical experts' workload in making accurate decisions and extract the most significant features and improve the diagnostic speed and accuracy using a supervised learning algorithm and linear support vector machines (SVM).

Agarwal et al. [14] suggested to build an automated, robust, intelligent and hybrid system for the early diagnosis and classifying of brain tumor the Auto Contrast Enhancer, Tumor Detector and Classifier to efficiently provide on-demand contrast improvement of poor contrast MRI images for the early diagnosis and classification of brain tumors.

Islam et al.[15] proposed for diagnosing brain tumors using deep learning and a meta-heuristic algorithm. The method involves extracting features from brain MRI images using AlexNet, reducing the complexity of AlexNet by employing an Extreme Learning Machine (ELM) network as a classification layer, and fine-tuning the parameters of the ELM network using an Amended Grasshopper Optimization Algorithm (AGOA).

Patil et al. [16] presented a technique for identifying brain cancers using digital replicas and advanced machine learning algorithms by analysing MRI images. Images obtained from MRI machines are stored in a centralised cloud using Internet of Things (IoT) digital devices. The input pictures and other health-related data are then retrieved from cloud storage. The Particle Swarm Optimization approach chooses features from images.

Zhu et al. [17] introduced a novel approach to brain tumor classification by exploring three pre-trained convolutional neural network (CNN) models: DenseNet201, EfficientNetB5, and InceptionResNetV2, combined with softmax activation for feature extraction.

Dang et al. [18] applied four transfer learning architectures- InceptionV3, VGG19, DenseNet121, and Mobile Net. We used a dataset with data from three benchmark databases of figshare, SARTAJ, and Br35H to validate the models.

Tseng et al. [19] optimized the eXtreme Gradient Boosting (XGBoost) approach through image processing and feature selection for the reliable identification of brain tumors. Images are improved using the Contrast-Limited Adaptive Histogram Equalization (CLAHE) method.

Sultanpure et al. [20] introduced the additional preprocessing techniques used in conjunction with the segmentation task, the combination of data augmentation and Window Setting Optimization was found to be the most effective tool for brain tumor classification.

3. RESEARCH PROBLEM

Due to manual interpretation the classification carried out on MRI images which are chaotic, time-consuming and erroneous as the infection number is rising beyond the limit. Moreover, the use of handmade attributes in conventional machine learning techniques limits the endurance and maximizes the cost of model. In some cases, the supervised learning techniques outperform unsupervised models, causing

over-fitting problems and being unsuitable for large-scale datasets. These issues emphasize the significance of designing a completely automated brain tumor classification system using deep learning models. The convolutional neural network (CNN) models are the most popular deep networks that are good at pattern recognition and classification. Generally, the neural network models are categorized into two types which include one-stage and two-stage approaches. The one-stage detection method is rapid but produces more false positives; on the contrary, the two-stage detection method offers high detection accuracy but is highly complex in terms of time. The existing brain Tumor Detection system implemented with Learning Rate (0.00001), large training time (1 to 3 hours to train for 10–20 epochs) and unsuitable to handle large dataset (255 MRI Images-2764 MRI Images). Existing brain tumor detection systems consumes more memory (6 GB to 12 GB of GPU memory). The classification accuracy of the model is greatly depends on the feature learning and extracting abilities of neural network.

4. RESEARCH OBJECTIVES

The objectives of the research work are:

- To increase the learning rate of the model and handle the large dataset a proposed CNN–KNN-based model with 10-fold cross-validation is introduced.
- To minimize the training time of the model a proposed Brain Neuro Net with HD Transformer method is introduced.
- To reduce the memory footprint of the model by using less memory size to store the weights of model which helps the model to be deployed with limited memory and processing power a proposed VITALT with soft-quantization approach is introduced.

5. RESEARCH METHODOLOGY

The research work consists of three main components: brain tumor classification using CNN-KNN, the Brain Neuro Net system, and VITALT system.

5.1 Detection of Brain Tumor in MRI Images using Deep Learning Method using proposed CNN-KNN

Create a deep learning model. The pipeline is broken down into various stages, starting with the receiving of raw data and ending with the system classification output. The following are the phases of the proposed system: Preparing the data,

Morphological Procedures, Segmentation of images, Extraction of Features and Categorization

5.1.1 Pre-Processing

A series of techniques applied to a scanned input image is known as pre-processing. Pre-processing separates the fascinating pattern 19 from the background by removing the undesirable data from the scanned images.

5.1.2 Morphological Procedures

Morphological Image Processing is a collection of non-linear image processing algorithms that deal with the shape or structure of image features.

5.1.3 Image Segmentation

Image segmentation is the process of dividing a picture into many parts in order to make it easier to analyze and interpret. Watershed segmentation was applied in this case. Image segmentation is a technique to split the image into multiple parts. The main objective of this separation is to facilitate the study and interpretation of images while preserving their quality. This technique is also used to trace the edges of objects in images. This technique labels pixels by their intensity and characteristics. These parts represent all the original images and acquire their characteristics such as intensity and likeness. Image segmentation helps create a 3D contour of the body for brain tumor detection. Segmentation is used in machine perception, malignant disease analysis, tissue volumes, anatomical and functional analyses, 3-D rendering techniques., virtual reality visualization and anomaly analysis, and object definition and detection.

5.1.4 Feature Extraction

Feature extraction is the process of reducing the number of requirements required to describe the actual data in an effective manner. The contour is drawn using a segmented picture.The algorithm analyses the relevance section by using the range values as the threshold, using the firmware evaluation method from a fragmented image.

5.1.5 Convolutional Neural Network

The two layers that comprise the convolution foundation are pooling and convolution shown in Figure 1.The images in the dataset are used to extract the

parameters. CNN classifiers are composed of completely linked layers that classify images based on attributes obtained from the convolution layer.

5.1.6 Acquiring Datasets

Throughout the training phase, the recognition system's performance was assessed and items were identified using the dataset. Brain tumor images are gathered into a database. A collection of images includes both well and ill people. We preserve all of the samples in the Joint Photographic Experts Group format.

5.1.7 Convolutional layer

The system has been programmed to be recognized by the network. Layer of convolution. CNN's initial layer gathers features from an input brain picture establishing the connection between pixels utilizing information about visual features and brain data. A kernel and an image matrix make up its two inputs. A convolution layer operates on an image to carry out tasks like edge recognition and filter blurring.

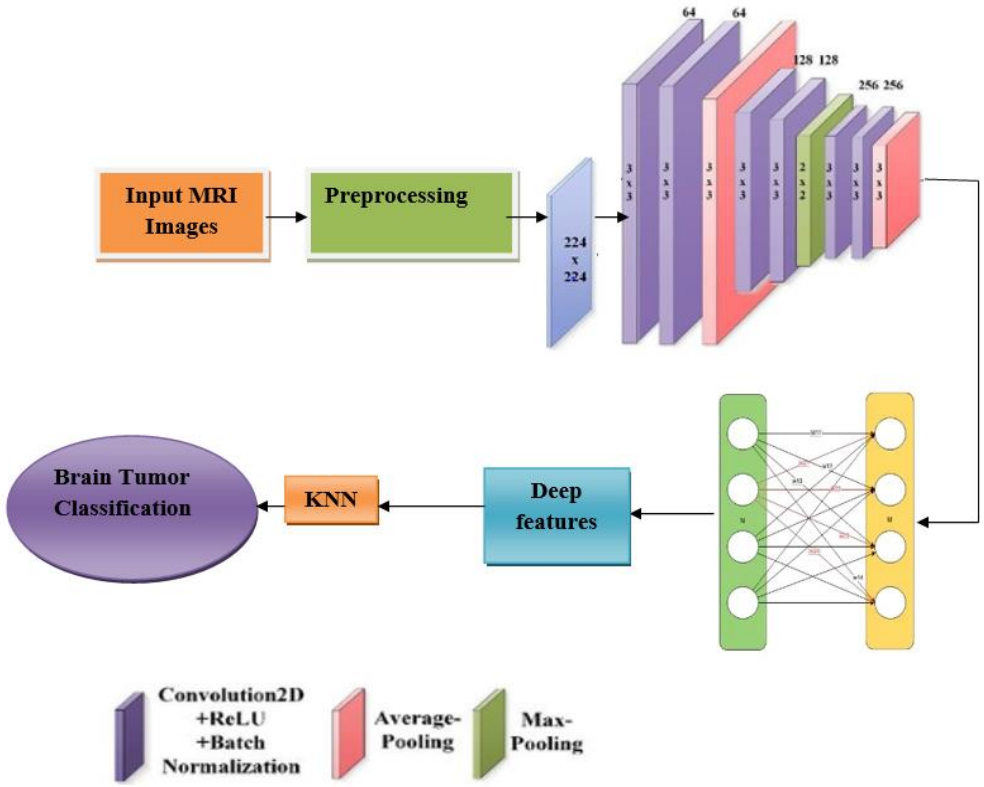

Figure 1: Image Classification technique for Brain Tumor Image classification using a proposed CNN-KNN

Following the convolution layer, the hidden layers generated the feature map. It is biased in the same way and has the same weight. The weights can be changed while the network is being trained.

5.1.8 Training phase

The data source is being used to train CNN. 90% of the images in the dataset have been used in the training phase, with the remaining images being taken for validation and testing. High resolution preprocessing is applied to the images, and during the learning process, an item of interest is chosen to help differentiate the brain tumor images from the surrounding ones. The size of the image was decreased to cut down on training time.

5.1.9 Image Preprocessing

During the learning process, an object of interest is chosen to help separate the brain tumor images from their surroundings. The photos are preprocessed with high definition. The image size was decreased in order to shorten the training period. Image augmentation is a method for growing the dataset. In the process of training, the network may pick up more features from the picture. The training dataset underwent a number of changes in order to expand the dataset. CNN was trained using the Image Net dataset and the perspective transformation. When a pre-trained model is implemented, it is possible that a new class pertinent to the diagnosis of brain tumor disease will be included in place of the original classifiers are Keep developing the unified model, Lift certain layers and leave the rest in place and Frost the convolution's base.

5.1.10 The Pooling Layer

Pooling layers are used to select brain tumor images from the lower half. It retrieves the picture from the backward convolution layer and shrinks the dataset while preserving the pertinent information regarding the section on brain tumors. Max-pooling chooses the largest element from the feature map for downsampling.

5.1.11 *Layer Completely Connected*

The max-pooling image is normalized into a vector before being added to a fully connected layer. A vector is generated by the feature matrix. All of the parts come

together to form a network model. Ultimately, the outputs were divided into two categories: healthy and sick brain images. During the testing phase, a representative image is provided to the convolution layer to be tested. The image is processed using a convolution layer. The image should be convoluted, and the matrix in the convolution layer should be activated with ReLU. Using its feature map, the testing sample brain images are coupled with the trained sample. The pooling layer receives the final convolution layer outcome, which is tangled by the outputs of numerous convolution layers. To lower the size of the dimensionality, pooling is used in the pooling layer.The output layer classifies brain images into healthy and ill groups based on the sigmoid activation function. The two classification techniques employed the sigmoid function. Therefore, we suggest using the softmax activation function. The softmax activation function has been used to categorize various disorders using brain imaging.

5.2 Brain Neuro Net : Advancing Brain Tumor Detection with Hierarchical Transformers and Multi Scale Attention

The ‘Brain Neuro Net is introduced for BT detection by using collaborative teacher-student networks such as Hierarchical DConv (HD) transformer and Multi Scale Attention (MSA) Network. The HD transformer model extracts the global features presented by employing MDHTA and SW-MDHTA. The MSA model extracts local features from the preprocessed images. This student network is constructed with dilation convolution because it has the ability to reduce the parameters. The extracted local and global features in the preprocessed images by using a student-teacher network are concatenated by employing a collaborative learning model. These concatenated features are trained by loss function for finding detection performances. The experiments are conducted with different measures and the comparative study provided that the Brain Neuro Net model achieved better detection performances of 98.63% related to other existing methods.

Brain Neuro Net is a Collaborative Learning Method with Teacher–Student Strategy for Brain Tumor Detection. Brain tumor refers to abnormal proliferation of brain tissues and the accurate detection is complex for radiologists. The accurate and early detection impact is based on the shapes intensities and location of the tumors. Therefore, a novel Brain Neuro Net approach is proposed for the detection of brain tumor using the collaborative teacher-student network. This proposed collaborative model is designed according to the combination of Hierarchical DConv (HD) transformer and

Multi Scale Attention (MSA) Network shown in Figure 2. The brain images are collected from Brain Tumor MRI and Br35H datasets and are preprocessed by using different processes such as image resizing, normalization and enhancement for image quality enhancement. The HD transformer model is taken as a teacher network that extracts global features in the preprocessed brain images. In this HD transformer model, the complex patterns are captured by asymmetric convolution and the computational complexity is decreased by shifting window technique. The MSA network is the student network that extracts local features in the preprocessed images and this model has a dilated convolutional block for decreasing parameters and retaining nodal information. The collaborative learning model is applied to concatenate both local and global extracted from the preprocessed images, and loss function trained the concatenated features for the prediction of brain tumors. The experimental validation is used to demonstrate the efficiency of the proposed model by using significant tumor detection-related performance measures. The comparative analysis showed that the proposed model attained a better detection accuracy of 98.63% rather to existing approaches.

5.2.1 Teacher network (HD transformer)

The modified form of a swim transformer is called as HD transformer which includes different operations such as patch partition, linear embedding, patch merging, and HD transformer block. The primary purpose of this HD transformer is global feature extraction from the MRI images.

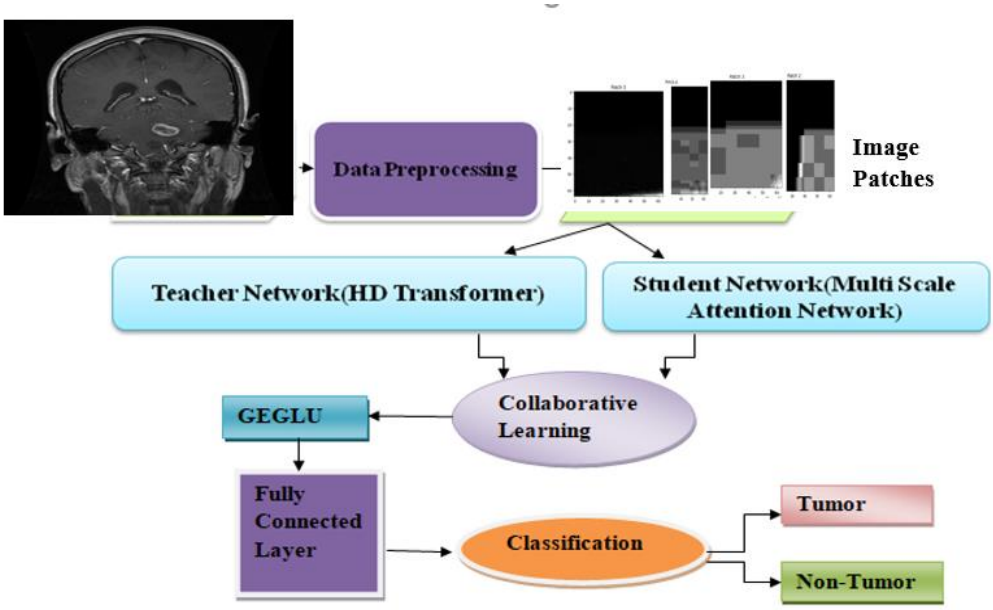

Figure 2: Architecture of Brain NeuroNet for Brain tumor Classification

5.2.1.1 HD transformer

At first, the images are cropped into different patches before entering the network. The linear embedding process converts each and every patch into one-dimensional vectors that are taken as input of the transformer model. The complex patterns are learned by asymmetric convolution with different filter sizes. For fostering inter-window communication, the W-MDHTA is followed by SW-MDHTA at each iteration. Layer norm is placed before each MDHTA module and the residual connections are after each MDHTA module shown in Figure 3. At last, the features are aggregated with residual learning operation by using an asymmetric convolution block. The complexity of feature maps is decreased through patch merging.

5.2.2 Student network (SN) (Multiscale attention (MSA) Network)

The SN is responsible for extraction of multi-scale local features that consist of a multi-scale attention network. This student network is designed based on the components of 1×1 convolution, 3×3 dilated convolution, LeakyReLU (Rectified linear unit), and Instance normalization (InstanceNorm). 1×1 convolution reduces the computational cost and the dilated convolution reduces the amount of parameters with increasing receptive field for reducing computations and retaining nodal information shown in Figure 4.The dilated convolution operation is performed with four different dilated rates.

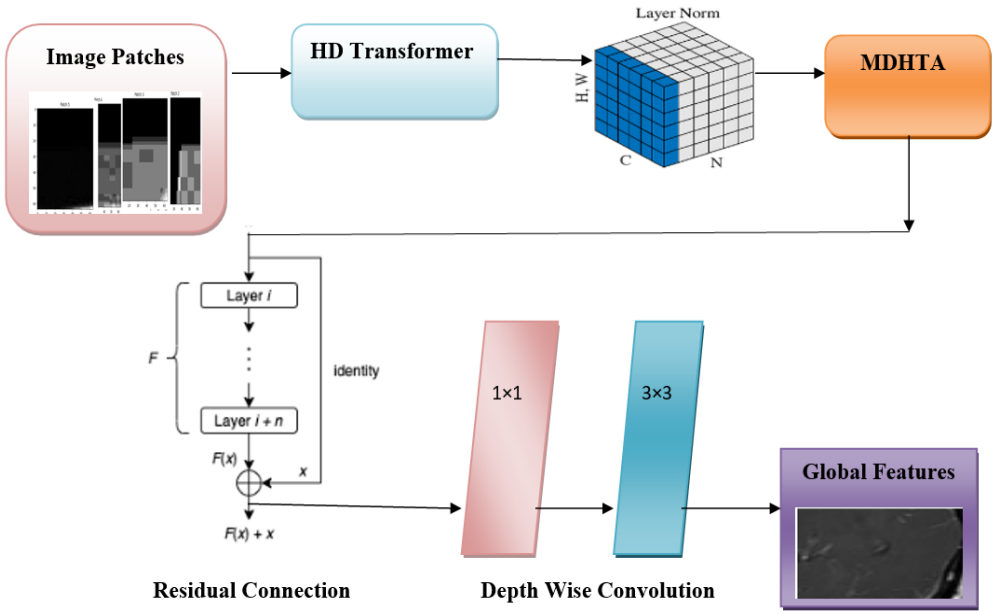

Figure 3: Architecture of Teacher Network in the proposed Brain Neuro Net

5.2.1.2 Multi-Dconv Head Transposed Attention

The calculation overhead is introduced by the transformers from the SA layer. The storage and processing time requirements for the dot product of queries with key interactions in the SA block grow quadratically. So, the MDHTA model is Brain Neuro Net with linear complexity in this teacher network to reduce these issues. The cross-variance over the channels is calculated to create an attention map (AM) for capturing the global condition. Moreover, the addition of depth-wise convolution is made to highlight local context that is performed before computing the covariance of the features to produce a global AM. For increasing local context, projections for the query, key, and value are derived from the layer-normalized tensor . The pixel-wise cross-channel contexts are aggregated by applying 1×1 and 3×3 which represents convolution and depth-wise convolution is applied after performing aggregation that encodes the channel-wise spatial condition.

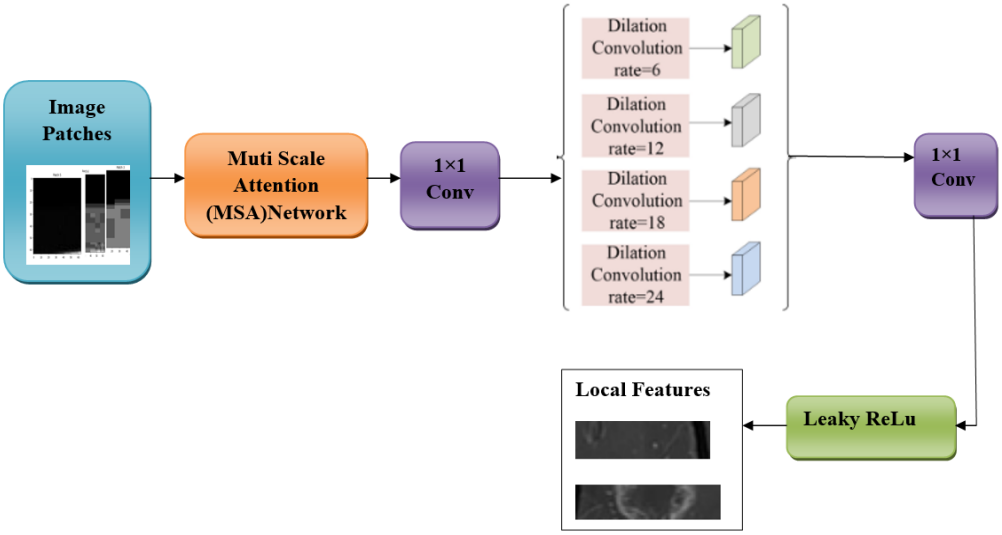

Figure 4: Architecture of Student Network in the proposed Brain Neuro Net

5.2.3 Collaborative learning

To speed up the training process, the features extracted from both teacher and student networks are fused by including cross-entropy loss. This collaborative strategy is used to train both transformer and MSA network in a cooperative reinforcement way for making two types of networks with its appropriate parameters. The first dense MLP layer is constructed by combining input layer, dense layer, and GeGLU activation, softmax output layer in dense block. The classical cross-entropy loss function (CCLF) is implemented in collaborative learning approach to optimize the features attained from both the HD transformer and MSA network. This loss

function utilizes image label and classification probability that provides the difference between image label and predicted results. The overall loss function is calculated based on the combination of classification loss and student-teacher loss functions that is evaluated by,

$$L_{over=} \lambda L_{TS}+(1-\lambda)L \quad (1)$$

The distillation rate is delineated as λ. Similarly, other feature extraction paths are encouraged to generate stronger feature representation.

5.3 VITALT: A Robust and Efficient Brain Tumor Detection System using Vision Transformer with Attention and Linear Transformation

Advanced brain tumor detection mechanism: A novel ViT with Attention and LTM (VITALT) based automated image classification mechanism is proposed to detect the presence of different classes of brain tumors in the MRI images.

Improvement of model Inference speed: The MRI images acquired from four different benchmark datasets have varied pixel resolutions that affect inference speed of model. For fair comparison across diverse image modalities, the preprocessing steps such as resizing, normalization and patch generation are executed that rescales and normalizes the images into uniform range as well divides each image into number of small square patches. This procedure increases the inference time and decreases training time of model.

Extraction of compact image representation: The image patches are embedded to form a high-dimensional vector representation using linear projection and non-linear activation function, making the model highly capable to learn more compact representations of MRI images.

Capturing local and global features: The ViT model with self-attention layer learns the relationship between the image patches to capture local and global features in the image. This helps to learn a more informative representation of the input image for the task of brain tumor classification.

Multi-dimensional feature fusion: The S-BiFPN model fuses the features across multiple scales to improve the accuracy of the model. It is important for the task of brain tumor classification, as tumors occur at different locations and scales in the brain. Enrichment of linear expression ability: The fused feature vectors are transformed using LTM module, which enhances the linear expression ability of the model, making it easier to understand the features that are most important for the task of brain tumor classification.

Minimizing memory footprints: The soft-quantization approach reduces the memory footprint of the model by using less memory size to store the weights of model. This helps the model to be deployed on resource-constrained devices with limited memory and processing power. Model weight adjustment using optimization algorithm: Backpropagation algorithm updates the weights of the model using the gradients of the loss function with respect to the model parameters. The optimal weights and biases provide high classification accuracy by minimizing the categorical cross-entropy loss function.

Quality refinement: The post-processing step which involves morphological operations such as erosion and dilation enhances classification accuracy by limiting the effects created on experimental phase of proposed brain tumor detection system. Quantitative results: The experimental results performed using different measures namely accuracy, precision, recall, f-score, error rate, dice score, Jaccard similarity index and Kappa coefficient shows that the proposed VITALT system offers useful insights to medical experts for taking effective treatment plans.

A fully automated detection model named 'VITALT' is proposed for effective and robust classification of brain tumor using MRI images. Figure 5 shows the overall architecture of the proposed VITALT system.

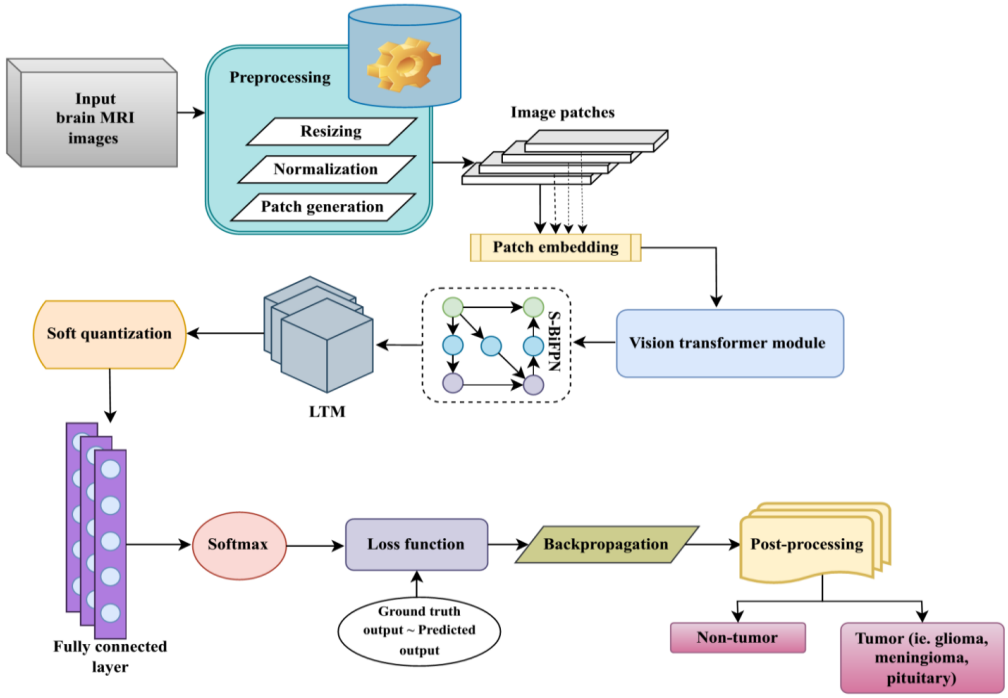

Figure 5: Overall architecture of the proposed VITALT system

The brain MRI images used for analysis are taken from four different benchmark datasets. The raw MRI image with data consistencies is resolved by preprocessing steps such as resizing, normalization and patch generation. The small square patches of images are embedded into a high-dimensional vector using a linear projection followed by a non-linear activation function. The embedded patches are trained through ViT with self attention module, S-BiFPN layer, LTM, soft quantization and head module one another to determine the presence or absence of brain tumor in the MRI images. The output of the network is post-processed to refine the predicted output and improve the accuracy of the model. A thorough description of each step involved in the proposed brain tumor detection system is elaborated in the following sub-sections.

5.3.1 Vision transformer model

The embedded patches are fed into the ViT model, which uses self-attention layers to capture global and local features of the image. The self-attention mechanism allows the model to attend to different parts of the input image and capture relationships between patches. The structure of ViT module is represented in Figure 6. The ViT model is based on the Transformer architecture, which was originally developed for natural language processing tasks. The ViT model applies the Transformer architecture to the task of image classification by treating the input image as a sequence of patches and applying self-attention layers to capture global and local features.

In the ViT model, each patch is represented by a learned embedding vector, which is fed into the self-attention layers. The self-attention mechanism allows the model to attend to different parts of the input image and capture relationships between patches.

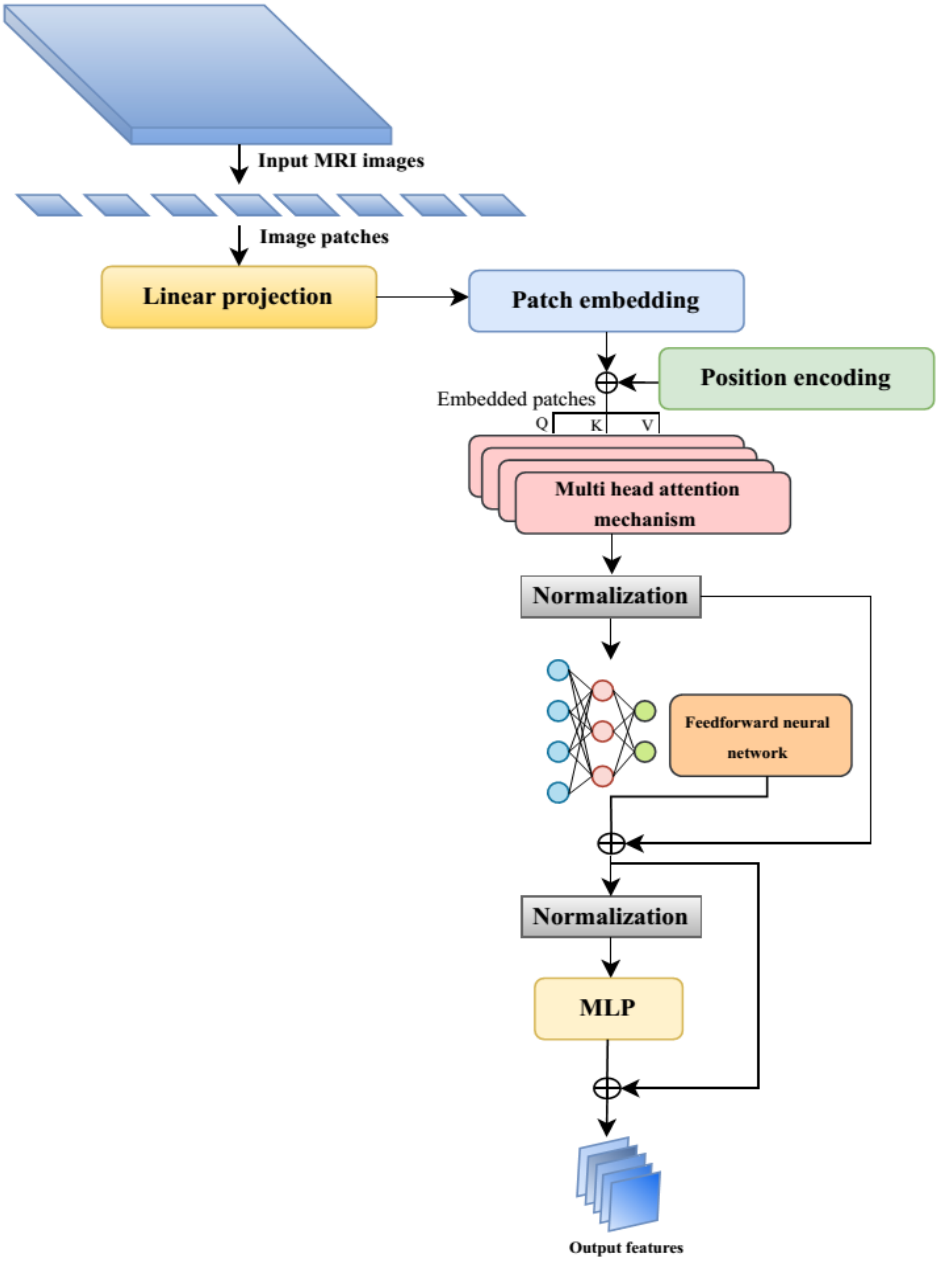

Figure 6: Structure of ViT module

Specifically, the self-attention mechanism computes a weighted sum of the embeddings for each patch, where the weights are learned during training and depend on the similarity between the embeddings.

4. CONCLUSION AND FUTURE WORK

This research work designed a robust and efficient brain tumor detection mechanism for accurate classification of tumor (ie. glioma, pituitary and meningioma) and non-tumor cases from MRI images. The pre-processing and post-processing steps are applied to MRI images to improve the inference speed and classification accuracy. Build a system for brain tumor detection that can distinguish between cancerous and

non-cancerous tumors involves several key components and technologies and provides a telehealth platforms to facilitate remote consultations and diagnostics.

5. THESIS ORGANIZATION

Chapter 1 – *"Introduction "*the effectiveness of machine learning and deep learning methods in classifying brain tumors based on medical imaging data. We discuss the leveraging Deep Learning techniques, to develop robust classification models.

Chapter 2 *"Literature Review"* provides comprehensive survey of different existing techniques used for deep learning based brain tumor detection model and highlighting the benefits and drawbacks of the existing methods that help to make the efficient model.

Chapter 3 "Detection of Brain Tumor in MRI Images using Deep Learning Method" proposed for enhance the classification rate and decrease the dependency of feature extraction and experimenter's expertise using a hybrid CNN–KNN-based model with 10-fold cross-validation.

Chapter 4 *"Brain Neuro Net: Advancing Brain Tumor Detection with Hierarchical Transformers and Multi Scale Attention"* proposed for handle overfitting issues the collaborative learning mode used to concatenate the local and global features that are collected from the networks of both teachers and students using Hierarchical DConv (HD) transformer and Multi Scale Attention (MSA) Network.

Chapter 5 "VITALT: A Robust and Efficient Brain Tumor Detection System using Vision Transformer with Attention and Linear Transformation*"* proposed to decreases training time of model the method

Chapter 6 *"Conclusion and Future work"* summarizes the key results, conclusions of this thesis and outlines the future scope of work.

6. REFERENCES

1. Ahamed, Md Faysal, Md Munawar Hossain, Md Nahiduzzaman, Md Rabiul Islam, Md Robiul Islam, Mominul Ahsan, and Julfikar Haider. "A review on brain tumor segmentation based on deep learning methods with federated learning techniques." *Computerized Medical Imaging and Graphics* (2023): 102313.

2. Atmakuri, Murali Krishna, A. Ganesh Ram, and V. V. K. D. V. Prasad. "Reliable image metrics-based brain tumor analysis using sensor deep learning technologies." *Measurement: Sensors* **29** (2023): 100864.

3. Turk, Omer, Davut Ozhan, Emrullah Acar, Tahir Cetin Akinci, and Musa Yilmaz. "Automatic detection of brain tumors with the aid of ensemble deep learning architectures and class activation map indicators by employing magnetic resonance images." *Zeitschrift für Medizinische Physik* (2022).

4. Talukder, Md Alamin, Md Manowarul Islam, Md Ashraf Uddin, Arnisha Akhter, Md Alamgir Jalil Pramanik, Sunil Aryal, Muhammad Ali Abdulllah Almoyad, Khondokar Fida Hasan, and Mohammad Ali Moni. "An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning." *Expert systems with applications* **230** (2023): 120534.

5. Patil, Suraj, and Dnyaneshwar Kirange. "Ensemble of deep learning models for brain tumor detection." *Procedia Computer Science* **218** (2023): 2468-2479..

6. Tahosin, Mst Sazia, Md Alif Sheakh, Taminul Islam, Rishalatun Jannat Lima, and Mahbuba Begum. "Optimizing brain tumor classification through feature selection and hyperparameter tuning in machine learning models." *Informatics in Medicine Unlocked* **43** (2023): 101414.

7. Turk, Omer, Davut Ozhan, Emrullah Acar, Tahir Cetin Akinci, and Musa Yilmaz. "Automatic detection of brain tumors with the aid of ensemble deep learning architectures and class activation map indicators by employing magnetic resonance images." *Zeitschrift für Medizinische Physik* (2022).

8. Puustinen, Sami, Hana Vrzáková, Joni Hyttinen, Tuomas Rauramaa, Pauli Fält, Markku Hauta-Kasari, Roman Bednarik et al. "Hyperspectral imaging in brain tumor surgery—evidence of machine learning-based performance." *World Neurosurgery* **175** (2023): e614-e635.

9. Raghuram, B., and Bhukya Hanumanthu. "Brain tumor image identification and classification on the internet of medical things using deep

learning." *Measurement: Sensors* **30** (2023): 100905.

10. Chatterjee, Surobhi, Rama Alkhaldi, Praggya Yaddav, Lakshmi Deepak Bethineedi, and Nourou Dine Adeniran Bankole. "Diagnostic performance of machine learning-based magnetic resonance algorithm vs conventional magnetic resonance imaging for predicting the likelihood of brain tumours." *Clinical Medicine*, **22 (**4) (2022),5-6.

11. Anantharajan, Shenbagarajan, Shenbagalakshmi Gunasekaran, Thavasi Subramanian, and R. Venkatesh. "MRI brain tumor detection using deep learning and machine learning approaches." *Measurement: Sensors* **31** (2024): 101026.

12. Ranjbarzadeh, Ramin, Payam Zarbakhsh, Annalina Caputo, Erfan Babaee Tirkolaee, and Malika Bendechache. "Brain tumor segmentation based on optimized convolutional neural network and improved chimp optimization algorithm." *Computers in Biology and Medicine* **168** (2024): 107723.

13. Lamba, Kamini, Shalli Rani, Monika Anand, and Lakshmana Phaneendra Maguluri. "An integrated deep learning and supervised learning approach for early detection of brain tumor using magnetic resonance imaging." *Healthcare Analytics* **5** (2024): 100336.

14. Agarwal, Monika, Geeta Rani, Ambeshwar Kumar, Pradeep Kumar, R. Manikandan, and Amir H. Gandomi. "Deep learning for enhanced brain Tumor Detection and classification." *Results in Engineering* **22** (2024): 102117.

15. Islam, Md Monirul, Prema Barua, Moshiur Rahman, Tanvir Ahammed, Laboni Akter, and Jia Uddin. "Transfer learning architectures with fine-tuning for brain tumor classification using magnetic resonance imaging." *Healthcare Analytics* **4** (2023): 100270.

16. Patil, Suraj, and Dnyaneshwar Kirange. "Ensemble of deep learning models for brain tumor detection." *Procedia Computer Science* **218** (2023): 2468-2479.

17. Zhu, Jing, Chuang Gu, Li Wei, Hanjuan Li, Rui Jiang, and Fatima Rashid Sheykhahmad. "Brain tumor recognition by an optimized deep network utilizing ammended grasshopper optimization." *Heliyon* **10**, no. 7 (2024).

18. Dang, Khiet, Toi Vo, Lua Ngo, and Huong Ha. "A deep learning framework integrating MRI image preprocessing methods for brain tumor segmentation and classification." *IBRO neuroscience reports* **13** (2022): 523-532.

19. Tseng, Cheng-Jui, and Changjiang Tang. "An optimized XGBoost

technique for accurate brain tumor detection using feature selection and image segmentation." *Healthcare Analytics* **4** (2023): 100217.

20. Sultanpure, Kavita A., Jayashri Bagade, Sunil L. Bangare, Manoj L. Bangare, Kalyan D. Bamane, and Abhijit J. Patankar. "Internet of things and deep learning based digital twins for diagnosis of brain tumor by analyzing MRI images." *Measurement: Sensors* **33** (2024): 101220.

LIST OF PUBLICATIONS

1. Poornam, S., and Saravanan Alagarsamy. "Detection of Brain Tumor in MRI Images using Deep Learning Method." In *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*, pp. 855-859. IEEE, 2022.

2. Poornam, S., and J. Jane Rubel Angelina. "BrainNeuroNet: advancing brain tumor detection with hierarchical transformers and multiscale attention." *International Journal of Information Technology* (2024): 1-8.

3. Poornam, S., and J. Jane Rubel Angelina. "VITALT: a robust and efficient brain tumor detection system using vision transformer with attention and linear transformation." *Neural Computing and Applications*, **36** (12) (2024), 6403-6419.